Exercise 1: Filter problems

Let $f(x_1, x_2|\boldsymbol{\mu})$ be the density function of the bivariate Normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{I}_2$. You are given the following data generating process (DGP):

- the target $Y \sim \texttt{Bernoulli}(0.5)$,
- the conditional density $p(x_1, x_2|Y = 1) = 0.5 \left(f(x_1, x_2|(1, -1)^\top) + f(x_1, x_2|(-1, 1)^\top)\right)$,
- the conditional density $p(x_1, x_2|Y = 0) = 0.5 \left(f(x_1, x_2|(1, 1)^\top) + f(x_1, x_2|(-1, -1)^\top)\right)$.

(a) Sketch the DGP

(b) Compute $\mathbb{P}(Y = 1|x_1 = \widetilde{x}_1), \mathbb{P}(Y = 1|x_2 = \widetilde{x}_2)$

(c) Compute $\mathbb{P}(Y = 1|x_1 = 1, x_2 = 1)$

(d) Explain what happens if we apply mutual information as filter in this scenario

Exercise 2: Filter simulation study

We want to implement a small simulation study:

- $d = 10$
- $n = 200$
- $\rho \in \{0, 0.1, 0.2\}$
- $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ with $\boldsymbol{\Sigma}_{ii} = 1.0$ and $\boldsymbol{\Sigma}_{ij} = \rho \quad \forall i \neq j$.
- $\boldsymbol{x}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}) \quad \forall i = 1, \ldots, n$
- $\boldsymbol{\beta} \in \mathbb{R}^d$ with $\boldsymbol{\beta}_{1:4} = 1, \boldsymbol{\beta}_{5:10} = 0$

We want to benchmark a linear model, a linear model with feature selection (with optimal number of features fixed), a linear model with feature selection (number of features is internally tuned). Repeat the benchmark 5 times.

(a) Implement the simulation study with `mlr3`
Hint: *https://mlr3book.mlr-org.com/chapters/chapter8/non-sequential_pipelines_and_tuning.html#sec-pipelines-featsel*, *https://mlr3.mlr-org.com/reference/benchmark.html#ref-examples*

(b) Analyze your findings

Exercise 3: Wrappers

You are given the following features and their respective BICs BIC_i with $i \in \{\{A\}, \{B\}, \{C\}, \{D\}, \{A, B\}, \{A, C\}, \{A, D\}, \{B, C\}, \{B, D\}, \{C, D\}, \{A, B, C\}, \{A, B, D\}, \{B, C, D\}, \{A, B, C, D\}\}$

Features	BIC_i
{A}	0.9
{B}	0.8
{C}	1.0
{D}	1.0
{A, B}	0.8
{A, C}	0.7
{A, D}	0.8
{B, C}	0.7
{B, D}	0.6
{C, D}	0.9
{A, B, C}	0.6
{A, B, D}	0.8
{B, C, D}	0.5
{A, B, C, D}	0.6

(a) Do forward search and note down each iteration.

(b) Do backward search and note down each iteration.